Broken Arrows

Teddy Furon, Patrick Bas

Hindawi Publishing Corporation
EURASIP Journal on Information Security
Volume 2008, Article ID 597040, 13 pages
doi:10.1155/2008/597040

Research Article

Broken Arrows

Teddy Furon[1] and Patrick Bas[2]

[1] *Project-Team TEMICS, INRIA Rennes-Bretagne Atlantique Research Centre, 35042 Rennes Cedex, France*
[2] *Gipsa-Lab Grenoble, CNRS, 38031 Grenoble Cedex, France*

Correspondence should be addressed to Teddy Furon, teddy.furon@inria.fr

Received 11 December 2007; Revised 18 April 2008; Accepted 21 August 2008

Recommended by Alessandro Piva

This paper makes an account of the design and investigations done for the still image watermarking technique used in the 2nd edition of the BOWS challenge. This technique is named "broken arrows" for some reasons given later on, and abbreviated "BA." This zero-bit algorithm is an implementation of a recent theoretical result by Merhav and Sabbag (2008) with precautions taken with respect to robustness, security, and imperceptibility. A new robustness criterion, based on the nearest border point of a cone, is proposed. The security constraint is taken into account by increasing the diversity of the watermark, sculpturing and randomizing the shape of the detection regions. The imperceptibility and robustness are also provided by adopting proportional embedding in the wavelet domain. The algorithm has been benchmarked using a database of 2000 images. For a probability of false alarm below $3 \cdot 10^{-6}$ and a PSNR of 43 dB, the overall robustness regarding various classical image processing seems a promising and strong basis for the challenge.

1. "BROKEN ARROWS" IN A NUTSHELL

1.1. Motivations

The watermarking technique "broken arrows" has been designed especially for the break our watermarking scheme 2nd edition (BOWS-2) contest. From the lessons learnt during BOWS-1, we had in mind to design a pure zero-bit watermarking scheme (no message decoding), which spreads the presence of the mark all over the host image. The BOWS-2 challenge is divided into three episodes with different contexts. The first episode aims at benchmarking the robustness of the technique against common image processing tools (compression, denoising, filtering, etc.). Thus, "BA" must be efficient so that it strongly multiplexes the original content and the watermarking signal in a nonreversible way when the secret key is not known. Moreover, no robustness against geometrical attacks is needed because they yield low PSNR values unacceptable in the contest. The second episode is dedicated to oracle attacks. The technique must be sufficiently simple so that the software implementation of the detector runs very fast because we expect a huge number of trials during this second episode. Counterattacks should be included if possible in the design. The third episode focuses on threats when many contents watermarked with the same secret key are released. The contenders are expected to deduce some knowledge about the secret key in order to better hack the pictures. "BA" must not be trivially hacked. This is not an easy task especially since zero-bit watermarking tends to lack diversity.

This technique is inspired from four articles of different fields: information theory [1], signal processing and game theory [2], statistics [3], and image processing [4]. During the design, we relied on the following key ideas.

(i) We do not know how to zero-bit watermark an image. However, the recent work [1] shows that the optimum scheme for Gaussian vectors under certain restrictions among which is the low complexity of the detector exactly matching our requirement.

(ii) Multiplicative embedding (aka proportional embedding) offers many advantages: an embedding that is easy to implement and compliant with the human visual system [4], plus a good approximation of game theoretical optimum solution for spread spectrum schemes [2].

(iii) One of the most difficult things in zero-bit watermarking is to assess that the false alarm probability is lower than a given level. Yet, one exception is

for detection regions shaped like hypercones where tractable numerical calculations exist [1, 3].

(iv) The wavelet domain is one of the best embedding domain even if the watermark signal has been created in another space, because it is compliant with the human visual system. There exists a fast wavelet transform based on the lifting scheme.

In a nutshell, the detection regions are represented by a set of slightly modified hypercones. The embedding is classically done by moving a feature vector $\mathbf{v}_X$ of the host content deep inside this detection region to obtain a watermarked vector $\mathbf{v}_Y = \mathbf{v}_X + \mathbf{v}_W$. The detection is performed by checking whether a feature vector $\mathbf{v}_Z$ extracted from a submitted image belongs or not to one of these hypercones (see Section 3).

1.2. Three general constraints

Other subtleties of BA are motivated by the general constraints in image watermarking, for example, security, robustness, and distortion.

Distortion

The visual distortion has been taken into account by choosing the medium and high frequencies of the image thanks to the wavelet transform (see Section 2.2) and applying the proportional embedding (see Section 4). The PSNR of the watermarked images is controlled during the embedding, resorting to norm conservation property of some orthogonal transforms (see Section 2) and by taking into account the proportional embedding step (see Section 4.1).

Robustness

BA relies on two techniques in order to have a decent robustness. The first one is commonly known as informed embedding. Vector $\mathbf{v}_Y$ is generated in order to be as far as possible from the border of the detection region (see Section 3.1.2). Furthermore, proportional embedding in a transform domain enables to merge two signals sharing the same statistical structure. The host is almost decorrelated in this transform domain like the watermark signal, while the watermark signal amplitude is shaped as the one of the host. This helps to be robust against denoising attacks like Wiener filtering.

Security

The original content is projected successively onto lower-dimensional subspaces in order to ease the creation of the watermark signal (see Sections 2.2 and 2.3). However, the first projection is private and depends on the secret key. This prevents the pirate from tracing the contents in the successive subspaces, and it restricts his playground to a very high-dimensional space. The dimension is almost as big as the number of pixels in the image. The detection region is composed of several regions introducing some diversity in the embedding because the host contents are pushed towards many different regions (see Section 3.2). We hope that this diversity brings some gain in security level in the sense that the private projection will remain secret even if many watermarked contents are observed. Finally, at the detection side, the security is also strengthened by randomizing the decision of the detector when the signal is near the frontier (see Section 5.1) and by introducing notches in the detection region (see Section 5.2).

2. FOUR NESTED SPACES

The embedding and the detection involve four nested spaces: the "pixel" space, the "wavelet" subspace, the "correlation" subspace, and (what we call) the "Miller, Cox & Bloom" plane (abbr. MCB plane). Index letters "X, Y, W" denote, respectively, the representatives of the original content, the watermarked content, and the watermark signal to be embedded. We use the following terminology and notations to denote the representatives in the different domains:

(i) "image" in the pixel space of width $W_{\mathbf{i}}$ and height $H_{\mathbf{i}}$: $\mathbf{i}_Y$, $\mathbf{i}_X$, $\mathbf{i}_W$,

(ii) "signal" in the wavelet subspace, which is a subset of $\mathbb{R}^{N_s}$: $\mathbf{s}_Y$, $\mathbf{s}_X$, $\mathbf{s}_W$ (due to the discrete nature of pixels values, this subspace and the following ones are not stricto sensu homomorphic to $\mathbb{R}^{N_s}$, $\mathbb{R}^{N_v}$, or $\mathbb{R}^2$),

(iii) "vector" in the correlation space, which is a subset $\mathbb{R}^{N_v}$: $\mathbf{v}_Y$, $\mathbf{v}_X$, $\mathbf{v}_W$,

(iv) "coordinates" in the MCB plane, which is a subset $\mathbb{R}^2$: $\mathbf{c}_Y$, $\mathbf{c}_X$, $\mathbf{c}_W$.

The diagram of the different processes necessary to obtain the different subspaces is depicted on Figure 1. The following subsections describe these subspaces and their specificities.

2.1. The pixel space

Images in this article are $H_{\mathbf{i}} \times W_{\mathbf{i}}$ matrices of 8-bit luminance values. We can always consider that the watermark in the pixel space is the difference between the watermarked image and the original image: $\mathbf{i}_W = \mathbf{i}_Y - \mathbf{i}_X$. This is not very useful, except that we impose a distortion constraint based on the PSNR, that is, a logarithmic scale of the mean square error between pixel values of images $\mathbf{i}_X$ and $\mathbf{i}_Y$:

$$\mathrm{PSNR} = 10 \log_{10} \frac{255^2}{\mathrm{MSE}}, \tag{1}$$

where MSE is the mean square error: $\mathrm{MSE} = (W_{\mathbf{i}} H_{\mathbf{i}})^{-1} \sum_{i=1}^{H_{\mathbf{i}}} \sum_{j=1}^{W_{\mathbf{i}}} \mathbf{i}_W(i,j)^2$, with $(W_{\mathbf{i}}, H_{\mathbf{i}})$ being the width and height of the image in pixels.

2.2. The wavelet subspace

As stated in the introduction, the wavelet transform is an excellent embedding domain because of its compliance to the human visual system.

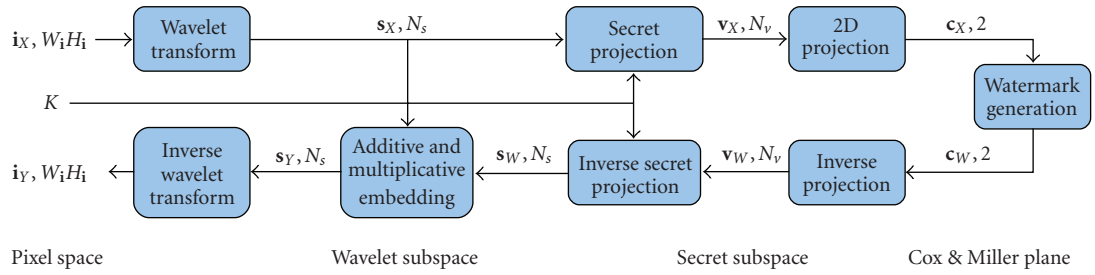

FIGURE 1: The different processes and entities involved in the BA embedding scheme. Each couple $(\mathbf{a}, N_a)$ represents the name and size of the vector $\mathbf{a}$.

We perform the 2D wavelet transform (Daubechies 9/7) on three levels of decomposition of the original image $\mathbf{i}_X$. This transform is very fast thanks to a very efficient lifted scheme implementation. We select the coefficients from all the bands except the low-frequency LL band. These $N_s = W_\mathbf{i}H_\mathbf{i}(1 - 1/64)$ wavelet coefficients are then stored into a signal $\mathbf{s}_X$ (a column vector) . In our implementation, the image dimensions must be multiple of 8. This signal lies in $\mathbb{R}^{N_s}$, a space we call the wavelet subspace. The low-low frequency band coefficients are kept in memory, and they will be used in the inverse extraction process.

The embedding process in this domain is in charge of mixing the host $\mathbf{s}_X$ and watermark $\mathbf{s}_W$ signals in a nonreversible way. The result is the watermarked signal $\mathbf{s}_Y = f(\mathbf{s}_X, \mathbf{s}_W)$. We mean by nonreversible the fact that an attacker observing $\mathbf{s}_Y$ should not be able to split it back to the two private signals.

The MSE in the wavelet subspace is equal to the MSE in the spatial domain because this wavelet transform conserves the Euclidean norm. Hence, to enforce the distortion constraint, we must have

$$||\mathbf{s}_Y - \mathbf{s}_X|| = 255\sqrt{W_\mathbf{i}H_\mathbf{i}} \cdot 10^{-\mathrm{PSNR}/20}. \tag{2}$$

2.3. The secret subspace

We use N_v secret binary antipodal carrier signals of size N_s: $\mathbf{s}_{C,j} \in \{-1/\sqrt{N_s}, 1/\sqrt{N_s}\}^{N_s}$. They are produced by a pseudorandom generator seeded by the secret key K. Their norm equals one, they are independent and we assume that they are orthogonal since their cross correlations are negligible (their expectations are zero, and their standard deviations equal $1/\sqrt{N_s}$) compared to their unitary norms when N_s is big. The host signal is projected onto these carrier signals: $v_X(j) = \mathbf{s}_{C,j}^T \mathbf{s}_X$. These N_v correlations are stored into a vector $\mathbf{v}_X = (v_X(1), \ldots, v_X(N_v))^T$. It means that $\mathbf{v}_X$ represents the host signal in the secret subspace. We can write this projection with the $N_s \times N_v$ matrix $\mathbf{S}_C$ whose columns are the carrier signals: $\mathbf{v}_X = \mathbf{S}_C^T \mathbf{s}_X$. The norm is conserved because the secret carriers are assumed to constitute a basis of the secret subspace: $\|\mathbf{v}_X\|^2 = \mathbf{s}_X^T \mathbf{S}_C \mathbf{S}_C^T \mathbf{s}_X \approx \|\mathbf{s}_X\|^2$.

The secret subspace has several advantages. Its dimension is much lower than the wavelet subspace; the vectors in this space are easier to manipulate. It brings robustness against noise or, in other words, it increases the signal to noise power ratio at the detection side, because the noise is not coherently projected onto the secret subspace. Moreover, it boils down the strong nonstationarity of the wavelet coefficients: components of $\mathbf{v}_X$ are almost independent and identically distributed as Gaussian random variables.

2.4. The Miller, Cox & Bloom plane

The MCB plane is the most convenient space because it enables a clear representation of the location of the hosts, the watermarked contents, and the detection boundary. This eases the explanation of the embedding and the detection processes. It is an adaptive subspace of the secret subspace whose dimension is two. We mean by adaptive the fact that this subspace strongly depends on the host vector $\mathbf{v}_X$. Denote $\mathbf{v}_c^\star \in \mathbb{R}^{N_v}$ as a secret vector in the secret subspace, with unitary norm. A basis of the MCB plane is given by $(\mathbf{v}_1, \mathbf{v}_2)$ such that

$$\mathbf{v}_1 = \mathbf{v}_c^\star, \qquad \mathbf{v}_2 = \frac{\mathbf{v}_X - (\mathbf{v}_X^T \mathbf{v}_1)\mathbf{v}_1}{\|\mathbf{v}_X - (\mathbf{v}_X^T \mathbf{v}_1)\mathbf{v}_1\|}. \tag{3}$$

Hence, the MCB plane is the plane containing $\mathbf{v}_C^\star$ and $\mathbf{v}_X$. As far as we know, [5] is the first paper proposing the idea that the watermark vector should belong to the plane containing the secret and the host, hence the name MCB plane.

The coordinates representing the host are $\mathbf{c}_X = (c_X(1), c_X(2))^T$ with $c_X(1) = \mathbf{v}_X^T \mathbf{v}_1$, and $c_X(2) = \mathbf{v}_X^T \mathbf{v}_2$. Note that whereas $c_X(2)$ is always positive, the sign of $c_X(1)$ is not a priori fixed. However, we will define $\mathbf{v}_c^\star$ so that $c_X(1)$ is indeed always positive (see (14)).

A useful property of the MCB plane is that the norm of the host vector is conserved. The denominator of $\mathbf{v}_2$ can be written as

$$\begin{aligned} ||\mathbf{v}_X - (\mathbf{v}_X^T \mathbf{v}_1)\mathbf{v}_1||^2 &= ||\mathbf{v}_X||^2 + c_X(1)^2 - 2\mathbf{v}_X^T(\mathbf{v}_X^T \mathbf{v}_1)\mathbf{v}_1 \\ &= ||\mathbf{v}_X||^2 - c_X(1)^2. \end{aligned} \tag{4}$$

Hence,

$$c_X(2)^2 = \frac{\left(\|\mathbf{v}_X\|^2 - c_X(1)^2\right)^2}{\|\mathbf{v}_X\|^2 - c_X(1)^2} = ||\mathbf{v}_X||^2 - c_X(1)^2, \tag{5}$$

so that $\|\mathbf{c}_X\|^2 = c_X(1)^2 + c_X(2)^2 = \|\mathbf{v}_X\|^2$. Now, the vector $\mathbf{v}_W$ to be added in the secret subspace is indeed first generated in the MCB plane, such that $\mathbf{v}_W = c_W(1)\mathbf{v}_1 + c_W(2)\mathbf{v}_2$. Then, $\|\mathbf{v}_W\|^2 = \|\mathbf{c}_W\|^2$.

3. EMBEDDING AND DETECTION

As mentioned in the previous section, the embedding first needs to go from the spatial domain to the MCB plane. Then, it creates the watermark signal and finally maps it back to the spatial domain. We have seen how to go from one subspace to another. We now explain the definition of the watermark representatives for the three domains.

We do not know what is the optimal way to watermark an image. This is mainly due to the nonstationarity of this kind of host. However, as mentioned earlier, host vectors in the secret subspace can be modeled as random white Gaussian vectors. We know what is the optimal way (in some sense) to watermark a Gaussian white vector according to [1]. The embedder has to create a watermarked vector as $\mathbf{v}_Y = a\mathbf{v}_X + b\mathbf{v}_c^\star$, where $\mathbf{v}_c^\star$ is a secret vector and a and b are scalars to be determined. This shows that the watermarked vector belongs to the plane $(\mathbf{v}_X, \mathbf{v}_c^\star)$, that is, the MCB plane. However, contrary to [1], we prefer to look for the optimum watermarked coordinate in the basis $(\mathbf{v}_1, \mathbf{v}_2)$ of the MCB plane.

3.1. The MCB plane

Knowing vector $\mathbf{v}_c^\star$, we perform the projection from the secret subspace to MCB plane as defined in (3). The detection region is defined by a cone of angle θ and abscissa direction $[0x)$ in the MCB plane such that $\mathbf{c}$ is considered watermarked if

$$\frac{|(1,0)\cdot\mathbf{c}|}{\|\mathbf{c}\|} = \frac{|c(1)|}{\|\mathbf{c}\|} > \cos(\theta). \tag{6}$$

The absolute value in the detection formula implies that the detection region is indeed a two-sheet cone as advised by Merhav and Sabbag [1].

The goal of the embedding process in this domain is to bring the coordinates of the watermarked vector $\mathbf{c}_Y$ deep inside the cone. There are actually several methods: maximizing a robustness criterion [6, Section 5.1.3], maximizing the detection score [1], or maximizing the error exponent [7]. These strategies assume different models of attack noise (resp., the noise vector is orthogonal to the MCB plane, the noise vector is null, or the noise vector is white and Gaussian distributed). We propose our own strategy which, in contrast, does not assume any model of attack as it foresees the worst possible noise. A geometrical interpretation makes the link between our strategy and the one from [6, Section 5.1.3].

3.1.1. Maximum robustness

This strategy is detailed in [6, Section 5.1.3]. Assume that $c_X(1) > 0$. We look for an angle $\tau \in [0, -\pi/2]$ which pushes the watermarked coordinates deep inside the detection region. This operation is defined by

$$\mathbf{c}_Y = \mathbf{c}_X + \mathbf{c}_W = \mathbf{c}_X + \rho\big(\cos(\tau), \sin(\tau)\big)^T. \tag{7}$$

The radius ρ is related to the embedding distortion constraint. We give its formula in Section 4.1.

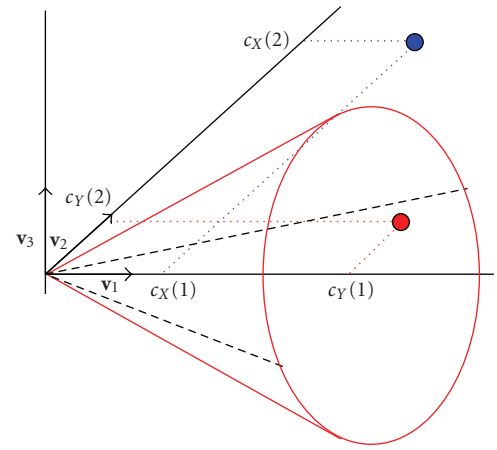

Figure 2: The hypercone and the MCB plane in the $(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3)$ basis.

Now, what does "deep inside the cone" mean? Cox et al. propose to maximize a robustness criterion defined by

$$R(\mathbf{c}_Y) = \max\big(0, c_Y(1)^2\tan(\theta)^2 - c_Y(2)^2\big). \tag{8}$$

Roughly speaking, R represents the amount of noise energy to go outside the detection region provided that $\mathbf{c}_Y$ is inside [6, Section 5.1.3]. The maximum robustness strategy selects the angle $\tau^\star$ maximizing the robustness: $\tau^\star = \arg\max_{\tau\in[0,-\pi/2]} R(\mathbf{c}_Y)$, where $\mathbf{c}_Y$ is a function of τ (7). This can be done via a dichotomy search or a Newton algorithm.

We would like to give a geometrical interpretation of this robustness criterion. Assume first that the attack noise is independent of the secret vector $\mathbf{v}_c^\star$ and of the host vector $\mathbf{v}_X$. Geometrically speaking, it means that this noise vector $\mathbf{v}_N$ is orthogonal to the MCB plane, giving birth to an orthogonal subspace spanned by $\mathbf{v}_3$ as depicted in Figure 2. A cut of the frontier of the detection region by the plane $(\mathbf{v}_2, \mathbf{v}_3)$ at the point $\mathbf{v}_Y$ shows a circle of radius $\mathbf{c}_Y(1)\tan(\theta)$. Figure 3 shows that the square norm of $\mathbf{v}_N$ needs to be at least equal to $(\mathbf{c}_Y(1)\tan(\theta))^2 - \mathbf{c}_Y(2)^2$ which is indeed $R(\mathbf{c}_Y)$.

3.1.2. A new criterion based on the nearest border point attack

The definition of the robustness explained above makes sense whenever the noise vector is orthogonal to the MCB plane. However, many attacks (filtering, compression, etc.) introduce a distortion which is indeed to be very dependent on the host vector. Hence, the previous assumption may not be realistic. We describe here a new embedding strategy maximizing the distance between the watermarked vector and the nearest border point on the detection region frontier. We first introduce it in an intuitive manner with geometrical arguments, and then we prove it with a Lagrange resolution which is indeed the best strategy.

Assume that the noise vector belongs to the MCB plane, then the shortest path to move outside the detection region is to push the watermarked vector orthogonal to the edge of the cone as shown in Figure 4. Hence, $\|\mathbf{c}_N\|^2 = (c_Y(1)\tan(\theta) - c_Y(2))^2\cos(\theta)^2$. It is very easy to show that

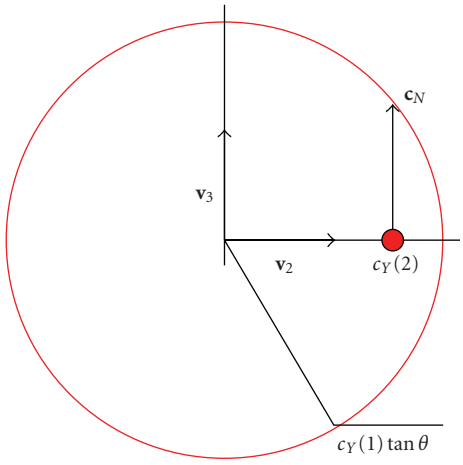

FIGURE 3: The minimal norm attack vector $\mathbf{c}_N$, when it is orthogonal to $\mathbf{v}_1$ and $\mathbf{v}_2$.

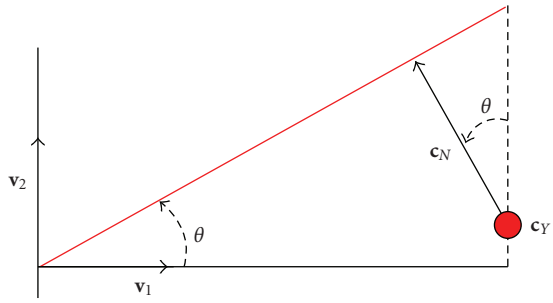

FIGURE 4: The border point attack in the MCB plane.

this norm is lower than $R(\mathbf{c}_Y)$. The embedding strategy should look for the coordinates $\mathbf{c}_Y$ maximizing the score $R'(\mathbf{c}_Y) = (c_Y(1)\tan(\theta) - c_Y(2))^2$ under the constraint that $\|\mathbf{c}_Y - \mathbf{c}_X\| = \rho$. In other words, we select the coordinates whose nearest border point attack needs the maximum noise energy. Intuitively, the embedder should push the host coordinates orthogonally to the edge of the cone so that $\mathbf{c}_Y = \mathbf{c}_X + \rho(\sin\theta, -\cos\theta)^T$. However, this intuition is wrong when the embedding circle $\|\mathbf{c}_Y - \mathbf{c}_X\| = \rho$ intersects the axis of the cone because there is no point in having a negative $\mathbf{c}_Y(2)$ which would decrease $R'(\mathbf{c}_Y)$. This detail is illustrated in Figure 5.

We now strengthen our rationale with a more rigorous approach. The noise vector can always be written as $\mathbf{v}_N = n_1\mathbf{v}_1 + n_2\mathbf{v}_2 + n_3\mathbf{v}_3$, where (n_1, n_2) are its coordinates in the MCB plane (the one defined by the original vector $\mathbf{v}_X$), and n_3 is the remaining component orthogonal to the MCB plane. Let us look for the nearest border point attack noise that is finding which point $\mathbf{v}_N^\star$ located on the cone minimizes the Euclidean distance $\|\mathbf{v}_N - \mathbf{v}_Y\|$ to a point $\mathbf{v}_Y = (y_1, y_2, 0)$ in the MCB plane and inside the cone. This question can be mathematically formulated by

$$(n_1, n_2, n_3)^\star = \arg \min_{n_1^2\tan(\theta)^2 = n_2^2 + n_3^2} \|\mathbf{v}_N - \mathbf{v}_Y\|. \quad (9)$$

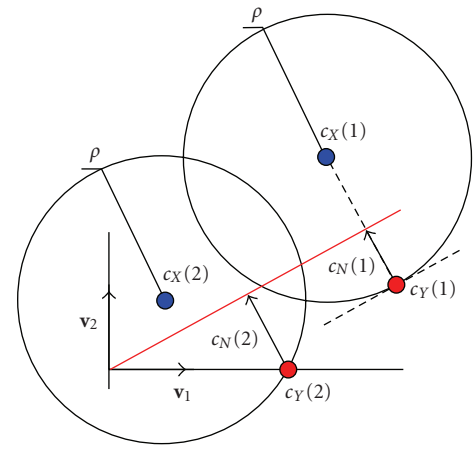

FIGURE 5: The two different embedding cases in the MCB plane.

A Lagrange resolution gives a point $\mathbf{v}_N^\star$ function of the coordinates of $\mathbf{v}_Y$, and the minimum distance $d_{\min}(\mathbf{v}_Y) = \min(\|\mathbf{v}_N^\star - \mathbf{v}_Y\|)$. Keep in mind that our real job is to place in the MCB plane coordinate $\mathbf{c}_Y$ at a distance ρ from $\mathbf{c}_X$, while maximizing this minimal distance:

$$\mathbf{c}_Y = \arg \max_{(y_1,y_2):(y_1-c_X(1))^2+(y_2-c_X(2))^2=\rho^2} d_{\min}(\mathbf{v}_Y). \quad (10)$$

This second constrained optimization is also easily solved by a Lagrange resolution. The study is divided into two parts depending on the first Lagrange resolution:

Case 1. If $y_2 > 0$, the minimal distance equals $d_{\min}(\mathbf{v}_Y) = y_1\sin(\theta) - y_2\cos(\theta)$ which is positive since $\mathbf{v}_Y$ is inside the cone. The nearest border point belongs to the MCB plane (i.e., $n_3^\star = 0$) with coordinates $(n_1^\star, n_2^\star) = (y_1, y_2) + (-\tan(\theta), 1)(y_1\tan(\theta) - y_2)\cos(\theta)^2$.

The second Lagrange resolution gives the watermarked coordinates: $\mathbf{c}_Y = \mathbf{c}_X + \rho(\sin(\theta), -\cos(\theta))$. Yet, this solution is acceptable only if $c_Y(2) > 0$, that is, $c_X(2) > \rho\cos(\theta)$. The maximum of the minimum square distance is then $\max(d_{\min}^2) = (c_X(2)\cos\theta - c_X(1)\sin\theta - \rho)^2$. Vector instances are shown in Figure 5 with superscript (1).

Case 2. If $y_2 = 0$ (i.e., $\mathbf{v}_Y$ is on the axis of the cone), then $d_{\min}(\mathbf{v}_Y) = y_1\sin(\theta)$, and the nearest border points are located on a circle: $n_1^\star = y_1\cos(\theta)^2$, and $n_2^{\star 2} + n_3^{\star 2} = n_1^{\star 2}\tan(\theta)^2$.

The distortion constraint allows to place the watermarked coordinates on the axis of the cone only if $c_X(2) < \rho$, and then $\max(d_{\min}^2) = (\sqrt{\rho^2 - c_X(2)^2} + c_X(1))^2\sin(\theta)^2$. We rediscover here the embedding proposed in [1, Theorem 2], where optimal parameters are given by [1, (33)]. This is the "erase" strategy where the embedder first erases the noncoherent projection of the host and then spends the remaining distortion budget to emits a signal in the direction of the secret vector $\mathbf{v}_c^\star$. Vector instances are shown in Figure 5 with superscript (2).

The two cases are possible and compete if $\rho\cos\theta \leq c_X(2) \leq \rho$. A development of the two expressions of $\max(d^2_{\min})$ shows that, in this case, the first case gives the real maximum minimum distance. Denote $R' = \max d^2_{\min}$. Our embedder amounts to place $\mathbf{c}_Y$ to maximize this criterion.

If $c_X(2) \leq \rho\cos\theta$, then

$$\begin{aligned}\mathbf{c}_Y &= \left(c_X(1) + \sqrt{\rho^2 - c_X(2)^2}, 0\right)^T,\\ R' &= \left(\sqrt{\rho^2 - c_X(2)^2} + c_X(1)\right)^2 \sin(\theta)^2,\\ \mathbf{c}_N &= c_Y(1)\sin(\theta)\left(-\sin(\theta), \cos(\theta)\right)^T.\end{aligned} \tag{11}$$

If $c_X(2) > \rho\cos(\theta)$, then

$$\begin{aligned}\mathbf{c}_Y &= \mathbf{c}_X + \rho\left(\sin(\theta), -\cos(\theta)\right)^T,\\ R' &= \left(c_X(1)\sin(\theta) - c_X(2)\cos(\theta) + \rho\right)^2,\\ \mathbf{c}_N &= \left(c_Y(1)\tan(\theta) - c_Y(2)\right)\cos(\theta)^2(-\tan(\theta), 1)^T,\end{aligned} \tag{12}$$

where $\mathbf{c}_N$ is the the nearest border point attacked coordinate.

In conclusion, a geometrical interpretation of the robustness criterion R given in [6, Section 5.1.3] gives us an idea of changing it for R'. This idea has been checked via a double Lagrange optimization. This new formula has links with the embedding strategy of [1] and also avoids the iterative search, as the optimal watermarked coordinates have now closed-form equation. The locus for the watermarked coordinates having the same robustness R' (aka contour of constant robustness) is the cone translated by the vector $\sqrt{R'}/\sin\theta\mathbf{v}_1$. This is quite a different constant robustness surface as the hyperbola (defined through (8)) gets very close to the border as the norm of the vector increases [5].

To go back to the wavelet subspace, we perform a double projection: firstly, we project $\mathbf{c}_W$ in the secret subspace and secondly, we project this result in the wavelet subspace to produce the watermark signal in the wavelet domain $\mathbf{s}_W$. Due to the use of orthonormal column vectors in both $\mathbf{S}_C$ (see Section 2.3) and $(\mathbf{v}_1, \mathbf{v}_2)$ matrices, this operation is defined by

$$\mathbf{s}_W = \mathbf{S}_C\left(c_W(1)\mathbf{v}_1 + c_W(2)\mathbf{v}_2\right) = \mathbf{S}_C(\mathbf{v}_1, \mathbf{v}_2)\mathbf{c}_W. \tag{13}$$

3.2. Increasing the diversity of the watermark signal

Zero-bit watermarking is known to provide weak security levels due to its lack of diversity. The detection region, for the moment, is only composed of one two-nappe hypercone around the axis supported by $\mathbf{v}_c^\star$. Analyzing several watermarked signals, an attacker might disclose the secret signal $\mathbf{S}_C\mathbf{v}_c^\star$ that parameterized the detection region, using clustering or principal component analysis tools [8–10].

For this security reason, we render the detection region more complex, defining it as the union of several two-nappe hypercones. In the secret subspace, we define a set $\mathcal{C}$ of secret directions with N_c secret unitary vectors: $\mathcal{C} = \{\mathbf{v}_{C,k}\}_{k=1}^{N_c}$. With the host signal being represented by $\mathbf{v}_X$ in this space, we look for the "nearest" secret direction from the host vector:

$$\mathbf{v}_C^\star = \operatorname{sign}(\mathbf{v}_C^T\mathbf{v}_X)\mathbf{v}_C, \quad \text{with } \mathbf{v}_C = \arg\max_{k\in\{1,\ldots,N_c\}} \left|\mathbf{v}_X^T\mathbf{v}_{C,k}\right|. \tag{14}$$

This secret vector is used for the embedding in the MCB plane. Chosen as is, projection $c_X(1)$ is always positive. At the detection side, the same vector has a high probability to be selected since the embedding increases correlation $\mathbf{v}_Y^T\mathbf{v}_C^{'}$.

We can predict two consequences. The first one is an advantage; we increase the probability of correct embedding. $\mathbf{v}_c^\star$ is chosen as the closest vector of $\mathcal{C}$ from $\mathbf{v}_X$, whence, for a given embedding distortion, it is more likely to push a vector $\mathbf{v}_Y$ in its hypercone. This acceptance region split into several areas mimics the informed coding used in positive rate or zero rate watermarking scheme such as dirty paper trellis or quantized index modulation. The second one is a drawback; the angle of the cones decreases with the number of cones in order to maintain a probability of false alarm below a given significance level. Consequently, narrower hypercones yield a lower robustness, as less attack distortion is needed to go outside.

The following subsections investigate this issue from a theoretical and an experimental point of view.

3.3. Modeling the host

In the wavelet subspace, one possible statistical model is to assume a Gaussian mixture. The wavelet coefficient $s_X(i)$ is Gaussian distributed but with its own variance σ_i^2: $s_X(i) \sim \mathcal{N}(0, \sigma_i^2)$. We will also pretend that they are conditionally independent given their variances, which is of course not exactly true [11]. In the secret subspace, the components of the host vector are Gaussian i.i.d. because the carrier signals are mutually independent, and the correlations are indeed linear combinations of Gaussian random variables: $v_X(j) \sim \mathcal{N}(0, \overline{\Sigma^2})$, with $\overline{\Sigma^2} = N_s^{-1}\sum_{i=1}^{N_s}\sigma_i^2$.

In the MCB plane, the statistical model is also very simple. Note first that $c_X(1)$ and $c_X(2)$ are not independent: $c_X(2) = \sqrt{c_X(1)^2 - \|\mathbf{v}_X\|^2}$. The first coordinate is defined as the maximum of N_c absolute values of correlation with unitary vectors. Hence, its cdf $F(c)$ is given by

$$\begin{aligned}F(c) &= \operatorname{Prob}(c_X(1) < c)\\ &= \prod_{\mathbf{v}\in\mathcal{C}} \operatorname{Prob}(|\mathbf{v}^T\mathbf{v}_X| < c)\\ &= \left(\Phi\left(\frac{c}{\sqrt{\overline{\Sigma^2}}}\right) - \Phi\left(\frac{-c}{\sqrt{\overline{\Sigma^2}}}\right)\right)^{N_c}\\ &= \left(2\Phi\left(\frac{c}{\sqrt{\overline{\Sigma^2}}}\right) - 1\right)^{N_c},\end{aligned} \tag{15}$$

where Φ is the standard normal cdf. As N_c becomes larger, the Fisher-Tippet theorem shows that $F(c)$ converges to the Gumbel distribution with a variance decreasing to zero with

a rate $1/\log N_c$ [12]. This allows us to roughly approximate $c_X(1)$ by its expectation which converges to the median value:

$$c_X(1) \sim \sqrt{\overline{\Sigma^2}}\,\Phi^{-1}\left(\frac{(1/2^{1/N_c}+1)}{2}\right), \tag{16}$$

where Φ^{-1} is the inverse normal cdf. By approximating also $\|\mathbf{v}_X\|^2$ by $N_v\overline{\Sigma^2}$, (5) shows the following ratio is approximately constant:

$$\frac{c_X(2)}{c_X(1)} \sim \sqrt{\frac{N_v}{\left[\Phi^{-1}((1/2^{1/N_c}+1)/2)\right]^2} - 1}. \tag{17}$$

Hence, the locus of the host coordinates in the MCB plane focuses more and more with N_c around a line passing by the origin and whose slope equals (17). Moreover, as mentioned above, the host coordinates get closer to the axis of the cone because the slope of the line is decreasing with N_c (see Figure 6).

3.4. Probability of false alarm

The hypercone is one of the very few detection regions where the probability of false alarm can be easily calculated provided that the host vectors pdf is radially symmetric, that is, only depending of the norm of the vectors. This is the case in BA, we can thus use work [3]. The probability that $\mathbf{v}_X$ falls inside a cone of angle θ is given by

$$\text{Prob}(\mathbf{v}_X \text{ in a two-nappe cone}) = \frac{I_{N_v-2}(\theta)}{I_{N_v-2}(\pi/2)}, \tag{18}$$

where $I_{N-2}(\theta)$ is the solid angle associated to angle θ in dimension N. We just bound P_{fa} with a classical union bound:

$$P_{fa} \leq N_c \frac{I_{N_v-2}(\theta)}{I_{N_v-2}(\pi/2)}. \tag{19}$$

As shown in Figure 6, the angle of the cone is moderately decreasing with N_c.

3.5. Experimental investigations

Figure 6 depicts the distributions of the coordinates of 5500 original images and their watermarked versions (PSNR of 45 dB) in their MCB plane for different numbers of cones while keeping the probability of false alarm below 10^{-6}. The host model is represented by the green line on the left. The bigger the number of cones is the better the approximative model feats the experimental distribution. The effect of the proposed strategy to maximize the robustness is clearly visible. The points representing watermarked contents are either located on the $\mathbf{v}_1$ axis or nearly distributed along the blue line on the right parallel to the line modeling the host coordinates. Host coordinates above the dotted blue line are just shifted by the vector $\rho(\sin\theta, -\cos\theta)^T$.

Figure 7 represents the robustness criterion R' calculated for these 5500 real host coordinates against $\|\mathbf{c}_X\|$ and enables to draw important remarks for constant embedding.

(i) For images with a low magnitude of $\|\mathbf{c}_X\|$, the robustness decreases with the number of cones.

(ii) For images with a high magnitude of $\|\mathbf{c}_X\|$, the robustness increases according to the number of cones.

(iii) For a given number of cones, there is a range of $\|\mathbf{c}_X\|$, for example, a class of images, where the robustness is maximal.

(iv) The average robustness is monotonically increasing with the number of cones. It tends to saturate for $N_c > 50$. This is an extremely surprising experimental result because we expected to have an optimal number of cones like the optimum number of codewords in Costa's theory [13].

This subsection has investigated the watermark embedding and detection only in the MCB plane. This exactly simulates an additive spread spectrum embedding where the watermark signal defined in (13) is directly added to the wavelet coefficients: $\mathbf{s}_Y = \mathbf{s}_X + \mathbf{s}_W$. This provides a tractable model in the MCB plane but it has many drawbacks, as we will see in the next section.

4. PROPORTIONAL EMBEDDING

The additive embedding has two major drawbacks. First, it does not comply with some psychovisual basics. The power of the watermark signal is constant all over the image, whereas the human eye is more sensitive on homogeneous regions than on textured regions and edges. Experimentally, the watermark appears as noise over uniform areas. The only way to avoid this artefact is to increase the PSNR, but then the robustness becomes weak. A second drawback is that additive embedding does not respect the power-spectrum condition [14] which states that the spectrum of the watermark has to be locally proportional to the spectrum of the host in order to be robust against denoising attacks. The intuition is that it is extremely hard or almost impossible to filter out the watermark signal if it shares the same statistical property than the host. A proportional embedding in the wavelet domain solves this two issues. The proportional embedding consists in locally adapting a gain prior the mixing: $\mathbf{s}_Y = \mathbf{s}_X + s_{W_p}$. The local gain is indeed proportional to the absolute value of the host wavelet coefficient:

$$s_{W_p}(i) = |s_X(i)|\, s_W(i). \tag{20}$$

In other words, the signal $\mathbf{s}_W$ is hidden in the content via a proportional embedding. Such an embedding in the wavelet domain provides a simple human visual system in the sense that it yields perceptually acceptable watermarked pictures for PSNR above 40 dB [4]. Moreover, this scheme has shown to be close to the optimal embedding strategy given by a game theory approach, but less computationally expensive [2]. However, some corrections are needed in the BA algorithm.

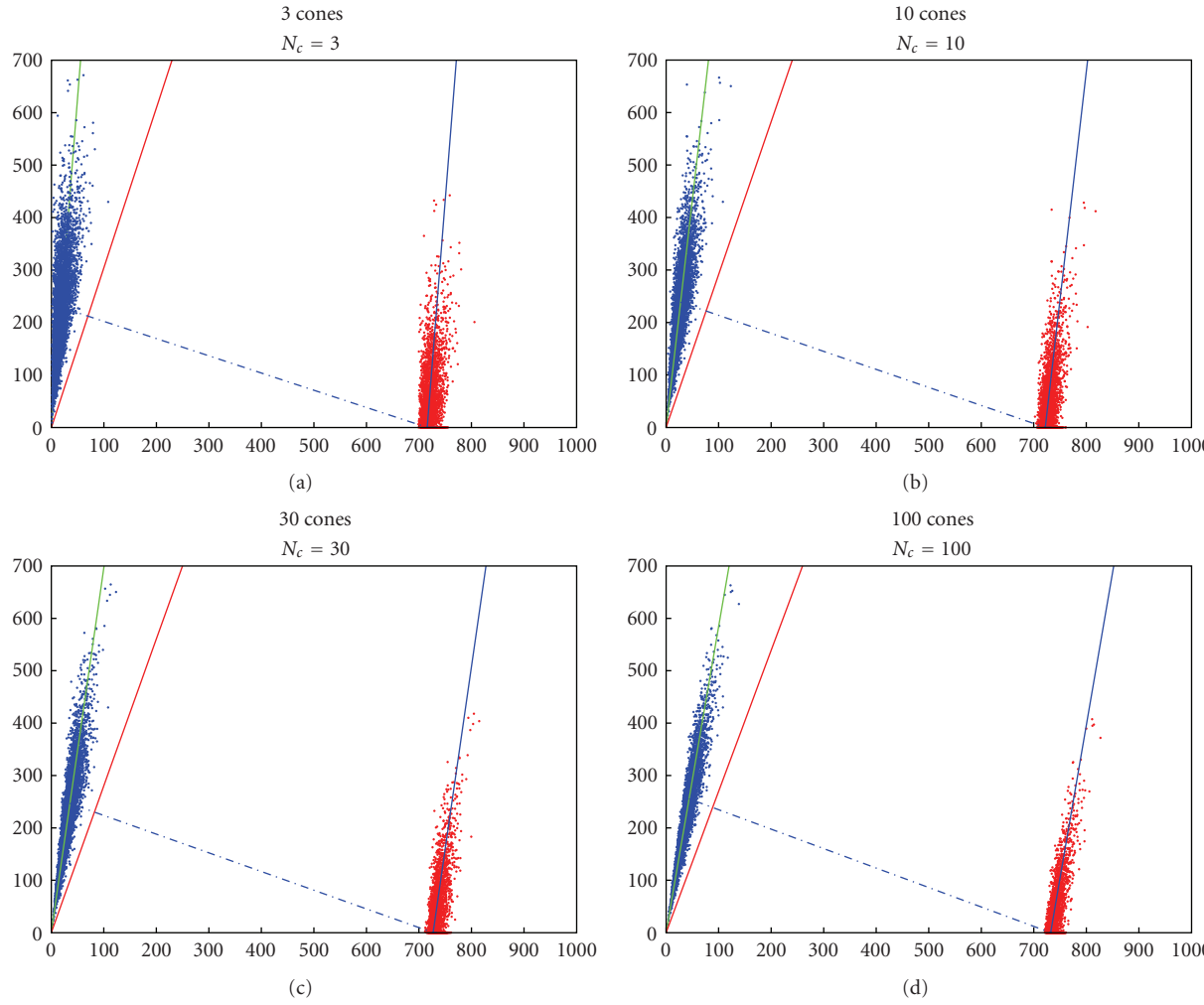

Figure 6: Distribution of the hosts (blue dots) and their watermarked coordinates (red dots) in their respective MCB plane. $P_{fa} = 10^{-6}$, equivalent PSNR = 45 dB.

4.1. Corrections

4.1.1. Impact on embedding distortion

The following equation links the norms of $\|\mathbf{s}_{W_p}\|$ and $\|\mathbf{s}_W\|$, assuming that $s_X(i)$ is independent from $s_W(i)$, $\forall i \in \{1,\ldots,N_s\}$:

$$\begin{aligned}\|\mathbf{s}_{W_p}\|^2 &= \sum_{i=1}^{N_s} |s_X(i)|^2 s_W(i)^2 \\ &\approx N_s^{-1} \sum_{i=1}^{N_s} s_X(i)^2 \sum_{i=1}^{N_s} s_W(i)^2 \\ &= \overline{S_x^2} \|\mathbf{s}_W\|^2,\end{aligned} \tag{21}$$

with $\overline{S_x^2} = N_s^{-1} \sum_{i=1}^{N_s} s_X(i)^2$. Hence, with (2), we must set the norm of $\mathbf{s}_W$ to

$$\|\mathbf{s}_W\| = \frac{255\sqrt{W_i H_i}}{\sqrt{\overline{S_X^2}}} 10^{-\mathrm{PSNR}/20}. \tag{22}$$

4.1.2. Equivalent projection

A difficulty stems from the fact that the proportional embedding is not a linear process. Assume that the embedder calculates watermarked coordinates $\mathbf{c}_Y$ in the MCB plane, and it mixes the corresponding watermark signal in the wavelet subspace with the proportional embedding. When the detector projects the watermarked signal back to the MCB plane, it does not find the same watermarked representative $\mathbf{c}_Y$. The watermarked signal is projected back onto the secret subspace in $\mathbf{v}_Y$ such that

$$v_Y(k) = v_X(k) + \sum_{i=1}^{N_s} \sum_{j=1}^{N_v} |s_X(i)| v_W(j) s_{C,j}(i) s_{C,k}(i). \tag{23}$$

We assume that the host wavelet coefficient $S_X(i)$ is statistically independent of the ith secret carriers samples in order to simplify this last expression in provided that $\mathbf{s}_{C,j}^T \mathbf{s}_{C,k} = \delta_{j,k}$:

$$v_Y(k) \approx v_X(k) + v_W(k) \overline{|S_X|}, \tag{24}$$

with $\overline{|S_X|} = N_s^{-1} \sum_{i=1}^{N_s} |s_X(i)|$.

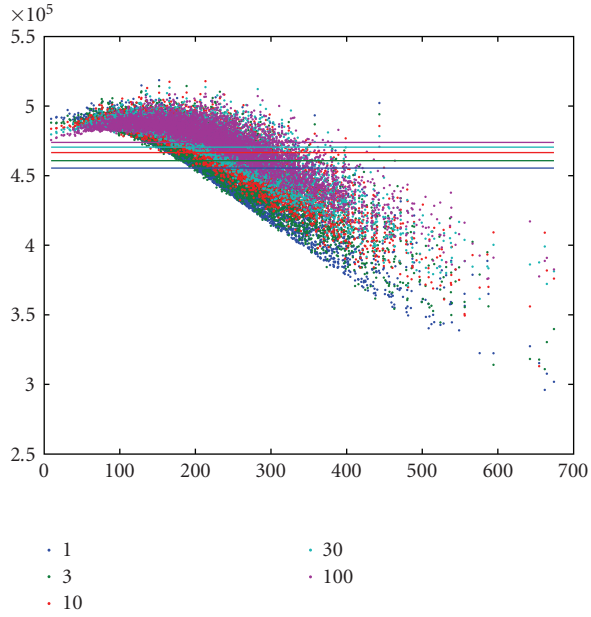

FIGURE 7: Computation of the robustness R' in function of $\|\mathbf{c}_X\|$ for different images and different number of cones. Average robustness is represented by horizontal lines.

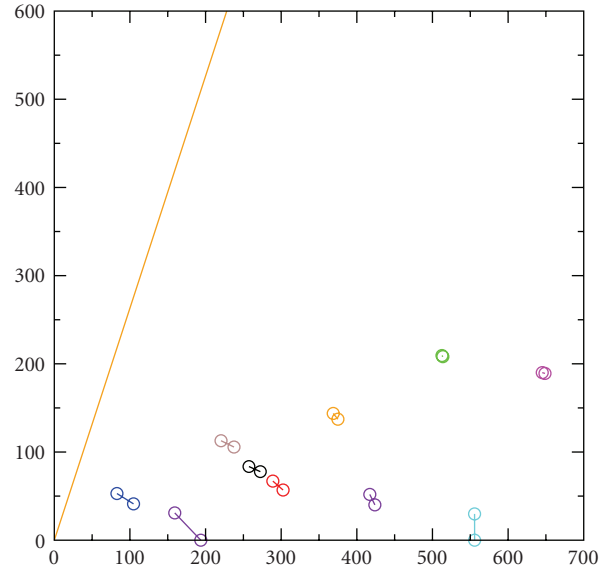

FIGURE 8: Deviations between the desired coordinates $\mathbf{c}_{W_p}$ at the embedding and the real coordinate $\mathbf{c}_Y$ retrieved at the detection. The coordinates are represented by circles as they appear in their own MCB plane, for 10 watermarked images.

At the embedding side, we take into account this phenomenon right in the MCB plane. We model it by searching the best watermark coordinates with a vector $\mathbf{c}_{W_p}$, which reflects the coordinates of the vector $\mathbf{c}_W$ after proportional embedding in the MCB plane. But, the coordinates to be projected back to the secret subspace is indeed $\mathbf{c}_W = \mathbf{c}_{W_p}/\overline{|S_X|}$.

These two corrections imply that even with a constant PSNR, the norm of $\mathbf{c}_{W_p}$ is different from a host image to another:

$$\|\mathbf{c}_{W_p}\| = \frac{\overline{|S_X|}}{\sqrt{\overline{S_X^2}}} 255\sqrt{W_i H_i} 10^{-\mathrm{PSNR}/20}. \tag{25}$$

In the MCB plane, the ratio $\overline{|S_X|}/\sqrt{\overline{S_X^2}}$ is the only difference between the additive and the proportional embedding methods.

4.2. Experimental investigations

We need to check that our model of proportional embedding in the MCB plane is actually working. Figure 8 shows ten couples of watermarked vectors in their own MCB plane. Each couple is composed of coordinate $\mathbf{c}_{W_p}$ used at the embedding and coordinate $\mathbf{c}_Y$ retrieved at the detection side when no attack occurred. There is a small difference, and the reason stems from all the approximations previously made: carriers are not orthogonal (Section 2.3), the embedding distortion is not exact (21), and projections in the secret subspace are modeled (24). The quantization in the spatial domain, after the inverse wavelet transform, is less disturbing. An iterative algorithm has already been proposed to better control the location of the watermarked coordinates [15, Section 3], but we believe that this difference is a second-order effect, and we prefer to keep the embedding process as simple as possible.

Other experimental works not described in this article showed us that bigger values of $\overline{|S_X|}$ and $\sqrt{\overline{S_X^2}}$ are expected when $\|\mathbf{v}_X\|$ is important, but there is almost no obvious statistical inference between $\overline{|S_X|}/\sqrt{\overline{S_X^2}}$ and the norm of the host vector. The expectation of this ratio is around 0.3 and 0.4, weakly increasing with $\|\mathbf{v}_X\|$. It has a strong variance around this expectation. The most important is that the ratio is always lower than 1. It means that embedding circle in the MCB plane is smaller with a proportional embedding than an additive embedding.

The final experimental work is a benchmark of four watermarking techniques. We used 2000 luminance images of size 512×512. These pictures represent natural and urban landscapes, people, or objects, taken with many different cameras from 2 to 5 millions of pixels.

The PSNR of the watermarked pictures is in average 42.7 dB. The visual distortion is invisible for almost all images. Figure 9 illustrates this with the reference image "Lena." A careful inspection shows some light ringing effects around the left part of the hat. However, there exist pictures where the embedding produces unacceptable distortion as shown in Figure 10. We explain this as follows. The common factor of these images is that they are composed of uniform areas (e.g., the sky) or textures with very low dynamic (e.g., the trees), and they have very few strong contours (the street lamp and the statue of Figure 10). Then, for a given distortion budget, the proportional embedding does not spread the watermark energy all over the image because most

Figure 9: The reference image "Lena" watermarked at PSNR = 42.6 dB.

Figure 10: One of the few images where the embedding provides a poor quality despite a PSNR of 42.7 dB. Ringing effect is visible around the statue and the street lamp.

wavelet coefficients are small, but it focuses the watermark energy on the very few strong wavelet coefficients. For the purpose of the challenge, we did not care of it but this drawback has to be improved for a real watermarking technique.

Four watermarking techniques with different embedding strategies have been benchmarked:

(i) maximization of the robustness criterion R defined by (8) with a proportional embedding,

(ii) maximization of the error exponent as detailed in [7] with a proportional embedding,

(iii) maximization of the new robustness criterion R' with a proportional embedding,

(iv) maximization of the new robustness criterion R' with an additive embedding.

We apply a set of 40 attacks mainly composed of combinations of JPEG and JPEG 2000 compressions at different

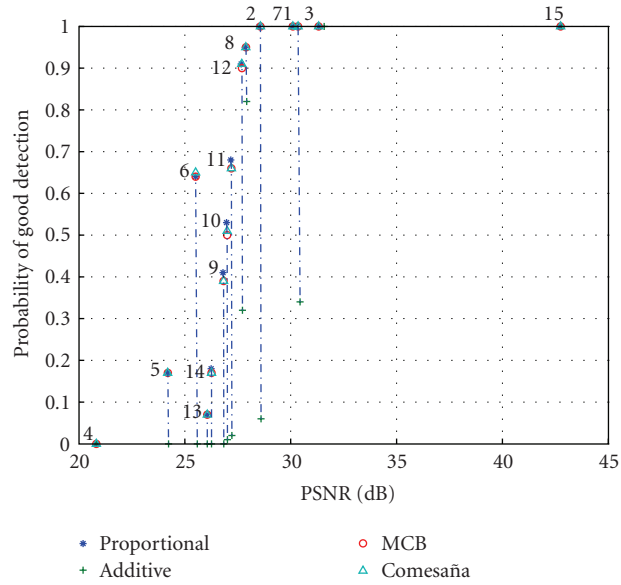

Figure 11: Probability of good detection versus average PSNR of attacked images for the four watermarking techniques: proportional embedding and new robustness criterion "$*$," additive embedding and new robustness criterion "+," proportional embedding and Miller, Cox & Bloom robustness criterion "$\circ$" [6, Section 5.1.3], proportion embedding with Comesaña's strategy "Δ" [7]. Selection of attacks: (1) denoise threshold 20, (2) denoise threshold 30, (3) JPEG $Q = 20$, (4) JPEG2000 $r = 0.001$, (5) JPEG2000 $r = 0.003$, (6) JPEG2000 $r = 0.005$, (7) scale 1/2, (8) scale 1/3 + JPEG $Q = 50$, (9) scale 1/3 + JPEG $Q = 50$, (10) scale 1/3 + JPEG $Q = 60$, (11) scale 1/3 + JPEG $Q = 70$, (12) scale 1/3 + JPEG $Q = 90$, (13) scale 1/4 + JPEG $Q = 70$, (14) scale 1/4 + JPEG $Q = 80$, (15) no attack.

quality factors, low-pass filtering, wavelet subband erasure, and a simple denoising algorithm. This latter consists in thresholding wavelet coefficients of 16 shifted versions of the image, afterward the inverse wavelet transforms are shifted back and averaged.

Figure 11 reports the impact of the 15 most significant attacks on the four techniques (the discarded attacks yield either lower PSNR or higher probabilities of detection). The probability of detecting the watermark (i.e., number of good detection divided by 2000) is plotted with respect to the average PSNR of the attacked images. Because these classical attacks produce almost the same average PSNR, the four points for a given attack are almost vertically aligned. Yet, the impact on the probability of detection is interesting; despite that the additive embedding allows bigger radius embedding circle in the MCB plane, this technique is the weakest. This stresses the fact that mixing signals with different statistical structures as for constant additive embedding is partly reversible. This is an Achilles' heel that even classical attacks take benefit of. Our embedding strategy gives average better results than the ones of Cox et al. (8) and of Comesaña et al. [7]. Yet, the improvement is really weak.

5. COUNTERATTACKS

In the BOWS-2 contest, the broken arrows algorithm has to face attacks linked to security. The first one is the oracle

attack whose goal is to disclose the shape of the detection region and/or to find nearest border point. The second one is based on information leakage, and the goal here is to try to estimate the secret subspace. We consequently decided to implement a counterattack in order to make these attacks (a bit) more complicated. The only solution we found to cope with information leakage attacks is to increase the diversity of the key by using several cones as explained previously in the paper (see Section 3.2). Initially, we also tried to increase the diversity of the key by using technique relying on perceptual hashing, but this technique was not mature enough to be implemented in the last final version of the algorithm. Regarding oracle attacks, we adopted three counterattacks presented below.

5.1. Randomized boundary

An attacker having unlimited access to the detector as a black sealed box can lead oracle attacks. Many of them are based on the concept of sensitivity, where the attacker tries to disclose the tangent hyperplane locally around a point (called sensitive vector) on the border of the detection region. A counterattack formalized by [16] is to slightly randomize the detection region for each call. This counterattack is very similar to the one that consists in having a chaotic boundary as proposed in [17, 18], both want to prevent an easy gradient ascent algorithm by making the detection border more difficult to analyse.

We process very simply by picking up a random threshold T uniformly distributed in the range $(\cos(\theta_{\max}), \cos((\theta_{\max} + \theta_{\min})/2)]$. $\theta_{\min}$ (resp., $\theta_{\max}$) has a corresponding probability of false alarm of $3\cdot10^{-7}$ (resp., $3\cdot10^{-6}$).

5.2. Snake traps

The "snake" is a new kind of oracle attack invented by Craver and Yu [19]. It consists in a random walk or a diffusion process in a constrained area of the space, which is indeed the detection region. This approach is a very efficient way to explore the detection region and to estimate parameters of the watermark detector. An important fact is that the snake tends to grow along the detection region border.

Our counterattack is to shape the boundary of the detection region, trapping the snakes in small regions to stop their growth. We draw "teeth" in the MCB plane, in the following way:

(i) if $|c_Y(1) - \Delta\lfloor c_{Y(1)}/\Delta \rfloor| < r$, then detection is positive if $c_Y(1) > \|\mathbf{c}_Y\|\cos(\theta_{\min})$,

(ii) else the watermark is detected if $c_Y(1) > \|\mathbf{c}_Y\|\cos(\theta)$, where θ is a random variable as explained above.

Δ and r set the periodicity and the width of the "teeth." Note that the teeth are longer as the vector is far away from the origin. Depending on the step of the random walk, we hope to increase the probability of trapping a snake as it grows. The teeth slightly reduce the size of the acceptance region, hence, the probability of false alarm is even lower.

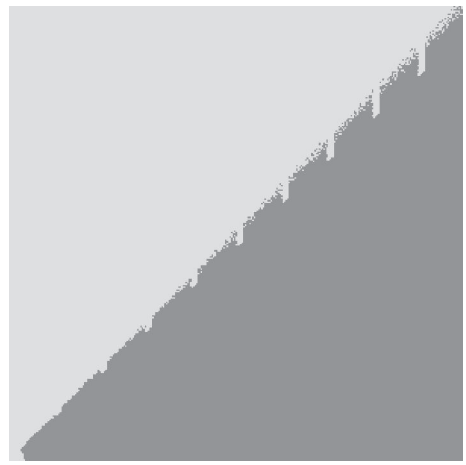

Figure 12: Detection results in the MCB plane. Dark-grey points represent contents detected as watermarked, light-grey points as not watermarked. The angle of the cone has been chosen in order to magnify the shape of the border.

Snakes almost grow infinitely in a cone because this detection region is not bounded (in practice, the pixel luminance dynamic bounds it). Hence, the average direction of several independent snakes can disclose the axis of the cone. Yet, we deal with several cones, and more importantly, the cones are indeed not disjoint for the considered probability of false alarm; the angle θ is always bigger than $\pi/4$ in Figure 6 and around 1.2154 rad in the final implementation. The snakes will then be trapped in a subspace of dimension N_c, where no average direction will emerge. This does not mean that snakes do no longer constitute a threat. A principal component analysis of several long snakes might disclose the secret subspace. We expect at least a strong increase of detection trials.

5.3. Camouflage of the cone

The detection score is virtually independent of a value-metric scaling. This is a nice robustness feature, but very few detectors provide this advantage. Hence, this leaves clues [19]. We consequently decided to conceal the use of hypercones by truncating it; a content is deemed not watermarked if $\|\mathbf{c}_Y\| < \lambda$. Note that the value λ has to be small enough to guaranty that the nearest border point is not located on the truncated section of the cone.

These three countermeasures result in a detection region depicted in Figure 12.

5.4. Nasty tricks

Concerning the challenge, we have the choice for the images proposed for the contest. We benchmark our watermarking technique over a set of 2000 images and against a bunch of common image processing attacks, in order to fine tune all the parameters, but also to investigate which images from this

database were the most robust. These latter ones are used for the first episode of the challenge. In the same way, we made a light JPEG compression to let participants think that the embedding domain is the DCT domain!

6. SOFTWARE IMPLEMENTATION

The BA software was developed in C using the libit [20] library in order to get fast embedding and detection schemes. During the whole contest, the embedding distortion is set by a targeted PSNR of 43 dB. In practice, due to pseudo-orthogonal carriers and the different approximations made in Section 4.1, the real PSNR is in between 42.5 dB and 43 dB.

A four-level wavelet decomposition is performed via libit with its very efficient implementation of a lifting step factorization using a Daubechies 9/7 biorthogonal wavelet [21]. The coefficients in subbands $\{HL_i, LH_i, HH_i\}; i \in \{1, 2, 3\}$ form the vector $\mathbf{s}_X$.

The pseudorandom generator is the Mersenne Twister pseudorandom number generator [22] whose seed, that is, the secret key K, is 128 bit long. The dimension of the secret subspace is $N_v = 256$. It is spanned by pseudo-orthogonal carriers on size $N_s = 258,048$. The Gram-Schmidt orthogonalization has been skipped because it is too much time-consuming. Antipodal carriers speed the correlation calculus because coefficients $\pm S_X(i)$ are accumulated in a sum which is in accordance with the sign of the corresponding carrier sample. We can also trade speed against memory; all the $N_s \cdot N_v$ carriers' samples are not stored in memory but they are generated as the need arises.

The number of cones N_c equals 30. There is no point in creating yet another set of secret directions for the axis of the cones. The secret subspace is already private via the secrecy of the antipodal carriers. Consequently, vector $\mathbf{v}_{C,k}$ is just the kth element of the canonical basis of the secret subspace. The angles $\theta_{\max}$ and $\theta_{\min}$ are chosen to obtain probabilities of false alarm lower than $3 \cdot 10^{-6}$ and $3 \cdot 10^{-7}$, respectively. These two probabilities bound the probability of false alarm of the whole system.

During the detection process, we choose a truncating parameter λ equal to 10, a period Δ equal to 30, and a width of the teeth r equal to 4.5. The random parameter to choose angle θ is computed using time as a seed of the pseudorandom generator.

For a 512×512 grey-scale image, the computational time for an embedding is of approximately 1.0 second for the embedding and 0.8 seconds for the detection on the BOWS-2 server (a 3-ghz Intel Xeon). Consequently, the BOWS-2 server, with 2 dual-core processors, has the possibility to detect around 350 000 images per day.

The source code of the BA embedding and detection schemes and the images used during the contest are available on http://bows2.gipsa-lab.inpg.fr.

7. CONCLUSION

The name "broken arrows" comes from the fact that the detection region is a set of cones shaped like heads of arrows, where the very end has been broken (see Section 5.3). Moreover, such a name suits perfectly the BOWS contest.

Designing a practical watermarking technique for a contest is a very challenging task. We would like to point out that a design is necessary done under constraints of time, man, and computer powers, with the sword of Damocles that a contender hacks the technique within the first hours of the challenge. Especially, the countermeasures presented in Section 5 have not been thoroughly tested due to lack of time. Consequently, a design is quite a different work than the writing of scientific paper. However, algorithms performing well in practice are often based on strong theoretical background. Not knowing the final results of the challenge by the time of writing, we humbly hope that lessons of scientific interest will be learnt.

ACKNOWLEDGMENTS

The authors thank Guillaume Stehlin and Francois Cayre for their knowhow in code optimization. They also thank Hervé Jégou, Vivien Chappelier, and Francois Cayre, the authors of the libit, for providing them such a simple and efficient C library as well as the denoiser software. The quality of this article has been really improved thanks to the careful reviewing of Cléo Baras, Francois Cayre, and the anonymous reviewers of EURASIP JIS. The authors also thank the French national ANR projects Nebbiano (ANR-06-SETI-009) and Estivale (ANR-05-RIAM-1902) for funding the BOWS-2 server. Last but not least, special thanks to Maj. Deakins and Capt. Hale for having supported them during the running of the challenge.

REFERENCES

[1] N. Merhav and E. Sabbag, "Optimal watermark embedding and detection strategies under limited detection resources," *IEEE Transactions on Information Theory*, vol. 54, no. 1, pp. 255–274, 2008.

[2] S. Pateux and G. L. Guelvouit, "Practical watermarking scheme based on wide spread spectrum and game theory," *Signal Processing: Image Communication*, vol. 18, no. 4, pp. 283–296, 2003.

[3] M. Miller and J. Bloom, "Computing the probability of false watermark detection," in *Proceedings of the 3rd International Workshop on Information Hiding (IH '99)*, A. Pfitzmann, Ed., vol. 1768 of *Lecture Notes in Computer Science*, pp. 146–158, Springer, Dresden, Germany, September 1999.

[4] F. Bartolini, M. Barni, V. Cappellini, and A. Piva, "Mask building for perceptually hiding frequency embedded watermarks," in *Proceedings of IEEE International Conference on Image Processing (ICIP '98)*, vol. 1, pp. 450–454, Chicago, Ill, USA, October 1998.

[5] M. L. Miller, I. J. Cox, and J. A. Bloom, "Informed embedding: exploiting image and detector information during watermark insertion," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '00)*, vol. 3, pp. 1–4, Vancouver, Canada, September 2000.

[6] I. Cox, M. Miller, and J. Bloom, *Digital Watermarking*, Morgan Kaufmann, San Francisco, Calif, USA, 2001.

[7] P. Comesaña, N. Merhav, and M. Barni, "Asymptotically optimum embedding strategy for one-bit watermarking under Gaussian attacks," in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents*, vol. 6819 of *Proceedings of SPIE*, pp. 1–12, San Jose, Calif, USA, January 2008.

[8] G. Doërr and J.-L. Dugelay, "Danger of low-dimensional watermarking subspaces," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP '04)*, vol. 3, pp. 93–96, Montreal, Canada, May 2004.

[9] F. Cayre, C. Fontaine, and T. Furon, "Watermarking security: theory and practice," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3976–3987, 2005.

[10] P. Bas and G. Doërr, "Practical security analysis of dirty paper trellis watermarking," in *Proceedings of the 9th International Workshop on Information Hiding (IH '07)*, T. Furon, F. Cayre, G. Doërr, and P. Bas, Eds., vol. 4567 of *Lecture Notes in Computer Science*, pp. 174–188, Saint Malo, France, June 2007.

[11] J. Liu and P. Moulin, "Information-theoretic analysis of interscale and intrascale dependencies between image wavelet coefficients," *IEEE Transactions on Image Processing*, vol. 10, no. 11, pp. 1647–1658, 2001.

[12] Wikipedia, "Fisher-tippett distribution—wikipedia, the free encyclopedia," August 2008.

[13] M. H. M. Costa, "Writing on dirty paper," *IEEE Transactions on Information Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[14] J. K. Su, J. J. Eggers, and B. Girod, "Analysis of digital watermarks subjected to optimum linear filtering and additive noise," *Signal Processing*, vol. 81, no. 6, pp. 1141–1175, 2001.

[15] G. L. Guelvouit and S. Pateux, "Wide spread spectrum watermarking with side information and interference cancellation," in *Security and Watermarking of Multimedia Contents V*, P. W. Wong and E. Delp, Eds., vol. 5020 of *Proceedings of SPIE*, pp. 278–289, Santa Clara, Calif, USA, January 2003.

[16] M. El Choubassi and P. Moulin, "On the fundamental tradeoff between watermark detection performance and robustness against sensitivity analysis attacks," in *Security, Steganography, and Watermarking of Multimedia Contents VIII*, E. J. Delp and P. W. Wong, Eds., vol. 6072 of *Proceedings of SPIE*, pp. 1–12, San Jose, Calif, USA, January 2006.

[17] J.-P. M. G. Linnartz and M. van Dijk, "Analysis of the sensitivity attack against electronic watermarks in images," in *Proceedings of the 2nd International Workshop on Information Hiding (IH '98)*, D. Aucsmith, Ed., vol. 1525 of *Lecture Notes in Computer Science*, pp. 258–272, Springer, Portland, Ore, USA, April 1998.

[18] M. Mansour and A. Tewfik, "Secure detection of public watermarks with fractal decision boundaries," in *Proceedings of the 11th European Signal Processing Conference (EUSIPCO '02)*, Toulouse, France, September 2002.

[19] S. Craver and J. Yu, "Reverse-engineering a detector with false alarms," in *Security, Steganography, and Watermarking of Multimedia Contents IX*, E. J. Delp and P. W. Wong, Eds., vol. 6505 of *Proceedings of SPIE*, pp. 1–10, San Jose, Calif, USA, January 2007.

[20] H. Jégou, V. Chappelier, and F. Cayre, "libit: Information theory and signal processing library," http://libit.sourceforge.net.

[21] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting steps," *Journal of Fourier Analysis and Applications*, vol. 4, no. 3, pp. 247–269, 1998.

[22] M. Matsumoto and T. Nishimura, "Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator," *ACM Transactions on Modeling and Computer Simulation*, vol. 8, no. 1, pp. 3–30, 1998.

EURASIP Journal on Wireless Communications and Networking

Special Issue on Wireless Physical Layer Security

Call for Papers

Security is a critical issue in multiuser wireless networks in which secure transmissions are becoming increasingly difficult to obtain in highly mobile and distributed environments. In his seminal works of the late 1940s, Shannon formalized the concepts of capacity (as a transmission efficiency measure) and equivocation (as a measure of secrecy). Together with Wyner's fundamental formulation of the wiretap channel in the 1970s, this work laid the groundwork for the area of wireless physical area security. Interest in this area has exploded in recent years, motivated by the rise of wireless networking in general and by the increasing interest in large mobile networks with light infrastructure, which are extremely difficult to secure by traditional methods.

The objective of this special issue (whose preparation is carried out under the auspices of the EC Network of Excellence in Wireless Communications NEWCOM++) is to gather recent advances in the area of wireless physical layer security from the theoretical, such as the analysis of the secrecy capacity of various channel models, to more practical interests such as the development of codes and other communication schemes that can provide security in real networks. Suitable topics for this special issue dedicated to physical layer security include but are not limited to:

- Opportunistic secrecy
- The wiretap channel with feedback
- Authentication over the wiretap channel
- Information theoretic secrecy of fading channels
- Secrecy through public discussion
- Wireless key distribution
- Multiuser channels with secrecy constraints
- MIMO wiretap channels
- Relay-eavesdropper channel
- Scheduling for secure communications
- Secure communication with jamming
- Game theoretic approaches for secrecy
- Codes for secure transmission
- Secure compression
- Cognitive approaches for secrecy
- Physical Secrecy and Common Randomness
- Secrecy with channel uncertainty

Before submission authors should carefully read over the journal's Author Guidelines, which are located at http://www.hindawi.com/journals/wcn/guidelines.html. Authors should follow the EURASIP Journal on Wireless Communications and Networking manuscript format described at the journal site http://www.hindawi.com/journals/ wcn/. Prospective authors should submit an electronic copy of their complete manuscript through the journal Manuscript Tracking System at http://mts.hindawi.com/, according to the following timetable:

Manuscript Due	December 1, 2008
First Round of Reviews	March 1, 2009
Publication Date	June 1, 2009

Guest Editors

Mérouane Debbah, Alcatel-Lucent Chair on Flexible Radio, Supélec, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France; merouane.debbah@supelec.fr

Hesham El-Gamal, Department of Electrical & Computer Engineering, Ohio State University, 205 Dreese Labs, 2015 Neil Avenue Columbus, OH 43210, USA; helgamal@ece.osu.edu

H. Vincent Poor, Department of Electrical Engineering, Princeton University, Engineering Quadrangle, Olden Street, Princeton, NJ 08544, USA; poor@princeton.edu

Shlomo Shamai, Department of Electrical Engineering, Technion, Technion City, Haifa 32000, Israel; sshlomo@ee.technion.ac.il